tions from the probable in the case of a correlated system of variables is such that it can be reasonably supposed to have arisen from random sampling. *Philosophical Magazine, Series 5* **50** 157–175.

Pearson, K. (1904). Mathematical contributions to the theory of evolution XIII: On the theory of contingency and its relation to association and normal correlation. *Draper's Company Research Memoirs. Biometric Series, No. 1.* [Reprinted in *Karl Pearson's Early Papers* (E. S. Pearson, ed.). Cambridge Univ. Press.]

Rubin, D. B. (1978). Bayesian inference for causal effects: The role of randomization. *Ann. Statist.* **6** 34–58.

Rubin, D. B. (1987). *Multiple Imputation for Nonresponse in Surveys.* Wiley, New York.

Schervish, M. J. (1987). A review of multivariate analysis (with discussion). *Statist. Sci.* **2** 396–433.

Simpson, E. H. (1951). The interpretation of interaction in contingency tables. *J. Roy. Statist. Soc. Ser. B* **13** 238–241.

Stigler, S. M. (1986). *The History of Statistics: The Measurement of Uncertainty before 1900.* Harvard Univ. Press.

Tanner, M. A. (1991). *Tools for Statistical Inference: Observed Data and Data Augmentation Methods. Lecture Notes in Statist.* **67**. Springer, New York.

Titterington, D. M., Smith, A. F. M. and Makov, U. E. (1985). *Statistical Analysis of Finite Mixture Distributions.* Wiley, New York.

Tuma, N. B. and Hannan, M. T. (1984). *Social Dynamics: Models and Methods.* Academic, Orlando, Fla.

Tuma, N. B., Hannan, M. T. and Groeneveld, L. P. (1979). Dynamic analysis of event histories. *American Journal of Sociology* **84** 820–854.

Yule, G. U. (1900). On the association of attributes in statistics. *Philos. Trans. Roy. Soc. London Ser. A* **194** 257–319.

Comment

David J. Bartholomew

This is an interesting and timely reminder of the important role which the social sciences have played, and continue to play, in the development of statistical methodology. I agree with so much of what the author says that it would be all too easy to make this contribution a repetition of the main points or a catalogue of additional supporting examples. Instead I wish to move the discussion in two other closely related directions—first by putting the emphasis on the inhibiting effect of the natural science influence on the development of statistical methodology and secondly by identifying current social science interests which place new demands on methodology.

I think the author is right in turning the spotlight on R. A. Fisher or, more exactly perhaps, the Fisherian tradition. Had it not been for Fisher's immense prestige the needs of social science might have continued to set an agenda for theoreticians as foreshadowed in the pioneering work of Quetelet and others. The core of modern statistical theory, centred on continuous variables, normal distributions, independence and additive models with the analysis of variance as its centrepiece has become the canon around which statistical education is built. The generalized linear model stands today as a fitting culmination of that tradition. The assumptions and the formulation of the models used are those required by the natural science problems which motivated Fisher and his followers and which still nourish much contemporary research. It is thus entirely understandable, though regretable, that the growth of multivariate analysis should, on the theoretical side, have been developed almost entirely around the multivariate normal distribution.

Perhaps the most striking example of this thesis is the laggardly way in which methods for categorical variables have become part of the statistician's portfolio. After the early excursions of Yule little note seems to have been taken of the fact that categorical variables are extremely common and, in a sense, more fundamental than their continuous counterparts. Until quite recently, measurement of association in two-way contingency tables was about as far as the education of most statisticians went. When they have been confronted with categorical data in practice they have had, for want of anything better, to force it into the Procrustean bed made for continuous variables by upgrading the level of measurement in more or less arbitrary ways. This is still very evident in the analysis of covariance structures where methods are developed for continuous variables and then adapted to categorical variables by the introduction of polychoric coefficients and such like. It is only now becoming apparent that there is a common structure underlying many such multivariate techniques which has been hidden by their diverse origins and notational idiosyncracies. To some

David J. Bartholomew is Professor of Statistics, Department of Statistical and Mathematical Sciences, London School of Economics and Political Sciences, Houghton Street, London WC2A 2AE, United Kingdom.

extent this has been fostered by those on the social science side. Lazarsfeld and Henry, for example, recognized the similarities between factor analysis and latent structure analysis but chose to emphasise the dissimilarities, which are secondary, at the expense of the similarities, which are not. The common basis of these methods which extends, for example, to multiple correspondence analysis has been sketched out in Bartholomew (1987) and could be paralleled elsewhere. Had the family of problems which these methods seek to tackle been approached from first principles rather than along existing tracks much needless effort might have been saved and social science might have progressed in better order.

A basic matter of great importance in social science, which is scarcely noticed in traditional statistics, is that of measurement. The typical statistical text introduces a topic by saying something like "Let x_1, x_2, x_3 . . . . be independent and identically distributed random variables" and takes that as the starting point. For the social scientist the real problem arises earlier with the question of how to measure the variable(s) involved. Professor Clogg rightly points out that social science variables can rarely be measured directly. Social theories are often couched in terms of quantities like quality of life or business sentiment whose measurement status itself is open to question. It is not simply a question having variables measured subject to error but of how to operationally define the variables in the first place. The ready availability of methods to analyse data has tended to divert attention away from the difficult part of the problem towards the relatively easy part. The complexity of an analysis is, unfortunately, more readily accepted as the yardstick of statistical sophistication than subtlety in the choice of variables.

The distributional assumptions of classical statistics are poorly suited to the analysis of data from social surveys. Even when simple random sampling is used the relevant theory for finite populations has often appeared as a somewhat awkward appendage to mainstream theory, and it often does not find a place in elementary texts. It is much more often the case, as the author says, that samples are stratified multi-stage cluster samples. The position is, perhaps, not quite as bad as Professor Clogg's brief and selective comment suggests. Work by Skinner, Holt and Smith (1989), for example, is one product of a vigorous research effort which is tackling the problems, but it only serves to underline the tardiness with which such efforts have been made. Again one is bound to speculate about the inhibiting effect of the natural science stranglehold on statistical thinking.

An equally pressing problem arises with data that are either not sampled at all or sampled in some poorly structured way. For want of anything better, standard statistical tests have often been used without much regard for what meaning they might have. There may well be good reasons for wanting to fit models to such data. Although one cannot speak meaningfully about the sampling errors of parameter estimates in such circumstances it still makes sense to be interested in their stability under variations in the data. Bootstrap-type calculations can be made in a routine fashion, but there appears to be no general guidance available. This is an area where current social science concerns are beginning to influence statistical methodology and where there is scope for a wider impact. This movement is particularly evident in one of the most notable omissions from Professor Clogg's survey. Gifi's (1990) substantial volume *Non-Linear Multivariate Analysis* comes from the Faculty of Social Sciences at the University of Leiden and amply demonstrates how different statistical theory might look if approached in a manner more in keeping with the character of the subject matter. Gifi also emphasises the visual exploration of data structures which is increasingly possible with modern computer graphics software. This is especially desirable with the large multivariate data sets to which social scientists frequently have access. It is noteworthy in this connection that the U.K. Economic and Social Research Council is making funds available for research on the analysis of large and complex data sets.

Of all the methods mentioned in the paper those which involve latent variables are, perhaps, most characteristic of social science. In origin they were almost entirely the work of social scientists, and even where similar problems arise elsewhere—as Clogg notes in the theory of finite mixtures and errors in variables—there is little cross-fertilization. Both sides have suffered as a result. For example, questions of statistical significance are rarely addressed in factor analysis, substantive "meaningfulness" often taking its place. The so-called axiom of local independence which the author mentions in relation to the LCM is also used in factor analysis and all similar latent variable models where it is scarcely noticed. We should remark in passing that it is not an assumption introduced to make the problem tractable but rather the test of whether all latent variables necessary to explain the interdependence of the manifest variables have been included. Although journals such as *Psychometrika* and the *British Journal of Statistical and Mathematical Psychology* (which surely merits a place in Clogg's list) show a marked resemblance to their mainstream counterparts I suspect the common readership is small. Generalizability theory, for example, is a term which has a conspicuous place in psychological statistics but would not be widely recognized by statisticians. In such cases one might hope that the influence of social science would not only be through the introduction of new methodology but that it would also stimulate statisticians to make improvements so that it can bear comparison with the best statistical practice.

These, and many examples like them, might well be held to support the proposition that although social science has produced a considerable amount of statistical methodology it has not sufficiently influenced the thought or direction of mainstream statistics. If this discussion helps to bring that about it will have been well worthwhile.

Comment: It's the Interplay That's Important

Paul W. Holland

While he hides it well behind the mask of scholarly indifference, Clogg is hopping mad. He has had it up to here with effete mathematical snobs (couldn't data analyze their way out of a paper bag) telling him that, at bottom, all the really good ideas, even in statistics, come from mathematics and the physical sciences. The last straw was reading the same bilge in, of all places, R. A. Fisher! An understandable fury, but two things that it is good for applied statisticians to develop are a tolerance of foolishness and a very thick skin.

I have been "doing statistics" in the social sciences for most of my life, but on occasion the opportunity to examine raw data from the physical sciences arises, and it has always struck me on these occasions just how familiar they appear, even though they hail from an allegedly distant part of the scientific landscape. At the forefront of research, the high signal-to-noise ratio that some associate with physical science data simply isn't there. If it is, then we aren't seeing the new stuff, just the well-established old stuff. What does seem to distinguish the social from the physical sciences is that, in the latter, progress is always being made in instrumentation and the noise levels eventually get lower. Signals eventually stand out with great precision, but when they do, it is old news; it's just Kuhn's "normal science."

Most social scientists get used to the fact that, generally speaking, the noise level does not decrease (increasing n just introduces new ways to slice up the data, and we are then often back to our noisy little samples again). In this respect, progress in the social sciences is hard to make. Learning to live with this fact is part of the early socialization of social scientists and of the statisticians who work with them. Most learn the hard way that a correlation over 0.8 is probably an error, and one exceeding 0.95 is an error for sure. (These limits must be modified somewhat for those who routinely correlate a variable with a slightly modified version of itself, but the same basic fact must be learned even there.)

The ubiquity of noise is why statisticians and statistics are useful to science—if there is no noise, no uncertainty, then there is no need for us or it.

Oh Statist! seek Unruly's feast,
and shun Perfection's meagre fare.

Human variability is a root source of the wide application of statistics to social, behavioral and medical science, and the lack of such a reliable source of noise is why statistical science has a more limited role in routine, empirical, physical science.

Aside from these comments, I have little to say about the endless and, to me, the totally sterile "social versus physical science" debate. I do have opinions about some of the points and examples that Clogg mentions in his interesting paper, so I'll concentrate the rest of my comments on them.

There is no question in my mind that real problems, based on real data, influence the development of statistical procedures. How can it be otherwise? Example: structural zeros in multi-way contingency tables. Categorical variables can have impossible combinations—for example, male hysterectomies. Ignoring the fact that there will never be a non-zero entry in such a cell can lead to a false impression of the association in such tables, but I doubt if anyone ever thought much about the implications of structural zeros until they saw them in real data. However, structural zeros have long appeared in continuous distributions, for example, the uniform distribution on the unit disk, but not much is usually made of this, essentially mathematical, example except to counter the proposition that a zero correlation implies independence. This is a continuous example of the "quasi-independence" that Clogg discusses because the two variables are as independent as they

Paul W. Holland is Distinguished Research Scholar at the Educational Testing Service and Fellow at the Center for Advanced Study in the Behavioral Sciences. Address correspondence to ETS, 21-T, Rosedale Road, Princeton, New Jersey 08541.